

REMARKS

Claims 1-37 are currently pending in the subject application and are presently under consideration. Claims 1 and 13-24 have been amended as shown at pages 2-5 of the Reply.

Applicants' representative thanks Examiner Colucci for the courtesies extended during the telephonic interviews conducted on October 10, 2007. Examiner was contacted to discuss the claim rejections under 35 U.S.C. §102(b) and 35 U.S.C. §103(a). During the interview a set of amendments were presented to overcome the rejections identified in the Office Action. These amendments have been incorporated into the claims as shown above. No agreement was reached regarding the amendments overcoming the cited art. Examiner indicated that further search and consideration was required.

Favorable reconsideration of the subject patent application is respectfully requested in view of the comments and amendments herein.

I. Rejection of Claims 1-4, 7, 8, 10, 14 and 16-22 Under 35 U.S.C. §102(b)

Claims 1-4, 7, 8, 10, 14 and 16-22 stand rejected under 35 U.S.C. §102(b) as being anticipated by Petkovic, *et al.* (US 6,185,527). It is respectfully submitted that this rejection should be withdrawn for at least the following reasons. Petkovic, *et al.* does not teach each and every element of the subject invention as recited in the subject claims.

For a prior art reference to anticipate, 35 U.S.C. §102 requires that "each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference." *In re Robertson*, 169 F.3d 743, 745, 49 USPQ2d 1949, 1950 (Fed. Cir. 1999) (*quoting Verdegaal Bros., Inc. v. Union Oil Co.*, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987)).

The subject claims relate to creating an audio thumbnail of an audio recording that is likely to allow a listener to easily identify the audio recording. Many times audio categorization systems provide for robust descriptive text related to songs. However, listeners many times are not able to recall the song solely based upon the text description and need to listen to the song to identify the song. Listening to the entire song is time consuming, but a representative audio thumbnail can allow the listener to identify the song by listening to a few seconds of the

thumbnail. For example, listeners are more likely to identify a song based upon a listening to a chorus or an instrumental section that repeats several times within the song. The repeating sections tend to be the sections of the song that are easily recalled from a listener's memory and associated with the song. In particular, independent claim 1 (and similarly independent claims 17 and 18) recites *producing one or more sets of clusters of fingerprints based upon all fingerprints within a set of clusters meeting a similarity threshold; and creating a thumbnail by employing a set of clusters that has at least two gaps between fingerprints, wherein a gap is a temporal space between two adjacent fingerprints that exceeds a predetermined threshold when fingerprints within a set of clusters are placed in sequential temporal order.*

Petkovic, *et al.* does not teach or suggest the aforementioned novel features as recited in the subject claims. The cited reference discloses a system for classification of audio content based upon identifying words in the audio content using various language based analysis techniques. However, the Petkovic, *et al.* does not teach grouping of similar fingerprints into sets of clusters based upon a threshold indicative of similarity. The cited reference discloses a threshold employed within a segment, not between two segments. Moreover, the cited reference fails to disclose identifying gaps exceeding a predetermined temporal threshold in sets of cluster in order to determine which set of clusters to employ in generating an audio thumbnail. The at least two gaps is indicative of a fingerprint that repeats at least three times, which is highly indicative of a repeating chorus of instrumental section. By employing similar fingerprints within a set of clusters and identifying a set of clusters that has at least two gaps of a threshold size, the subject claims can identify a chorus or instrumental section that repeats. Using fingerprints from this set of clusters allows for producing an audio thumbnail that will be more recognizable to a listener. Therefore, Petkovic, *et al.* fails to teach or suggest producing one or more sets of clusters of fingerprints based upon all fingerprints within a set of clusters meeting a similarity threshold; and creating a thumbnail from a set of clusters that has at least two gaps between fingerprints, wherein a gap is a temporal space between two adjacent fingerprints that exceeds a predetermined threshold when fingerprints within a set of clusters are placed in sequential temporal order.

Additionally, claim 14 recites *the heuristic component selects the set of clusters that has the highest value for the sum of the square of the mean spectral quality value determined for the set of clusters and the cluster spread quality value determined for the set of clusters.*

Petkovic, *et al.* is silent regarding a cluster spread quality value, as well as, the sum of the square of the mean spectral quality value determined for the set of clusters and the cluster spread quality value determined for the set of clusters, and fails to employ this summed value in determining the optimal set of clusters to employ in generating the audio thumbnail.

Moreover, claim 21 recites *the similarity threshold chosen adaptively based upon the audio file and used to help determine if two fingerprints belong to the same cluster set*. The cited reference discloses a threshold within a segment, but fails to disclose a threshold that is employed between two segments to determine if they are similar. Furthermore, Petkovic, *et al.* fails to disclose that the threshold is adaptive chosen based upon the audio file.

In view of the foregoing, applicants' representative respectfully submits that Petkovic, *et al.* fails to teach or suggest all limitations of independent claims 1, 17, and 18 (and claims 2-4, 7, 8, 10, 14 and 19-22 that depend there from), and thus fails to anticipate the subject claims. Accordingly, withdrawal of this rejection is respectfully requested.

II. Rejection of Claims 5, 9, 11, 12, 23, 24 and 34 Under 35 U.S.C. §103(a)

Claims 5, 9, 11, 12, 23, 24 and 34 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Petkovic, et al. (US 6,185,527) in view of Wells, et al. (US 2003/0086341 A1). It is respectfully submitted that this rejection should be withdrawn for at least the following reasons. Petkovic, *et al.* and Wells, *et al.*, alone or in combination, do not teach each and every element of applicants' invention as recited in the subject claims.

[T]he prior art reference (or references when combined) must teach or suggest all claim limitations. *See* MPEP §706.02(j). The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art and not based on applicant's disclosure. *See In re Vaeck*, 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991).

Claim 5, 9, 11, 12, 23, 24 and 34 depend from independent claims 1 and 18 respectively. As noted *supra*, Petkovic, *et al.* does not teach or suggest each and every element of the subject invention as recited in this independent claim, and Wells, *et al.* fails to make up for the aforementioned deficiencies of Petkovic, *et al.* Wells, *et al.* discloses a system for creating a fingerprint for an audio file that overcomes issues related to variations in content produced by

copying and using different encoding methods. The cited reference is concerned with creating as small as possible a fingerprint, where the fingerprint employs a plurality of audio file identifiers that help identify the audio file in order to provide redundancy in the identification. However, the cited reference fails to teach or suggest *producing one or more sets of clusters of fingerprints based upon all fingerprints within a set of clusters meeting a similarity threshold; and creating a thumbnail from a set of clusters that has at least two gaps between fingerprints, wherein a gap is a temporal space between two adjacent fingerprints that exceeds a predetermined threshold when fingerprints within a set of clusters are placed in sequential temporal order.* Wells, *et al.*, like Petkovic, *et al.*, is silent regarding a set of clusters containing similar fingerprints, as well as identifying at least two gaps of a given threshold size in determining portions of the audio file to employ in generating a fingerprint.

Furthermore, claim 23 recites *determining a parameter (D) describing how evenly spread clusters are, temporally, throughout an audio file.* As conceded in the Office Action, Petkovic, *et al.* fails to teach this novel feature. Contrary to assertions in the Office Action, Wells, *et al.* also fails to teach this novel feature. The Office Action cites paragraph [0231] of Wells, *et al.* as teaching this feature. However, the cited paragraph discloses an observed spread of fingerprint feature vectors across all fingerprints. This is a spread between fingerprints of different audio files that is used to determine the search (SRR) threshold for matching fingerprints between audio files. It is not a temporal spread of clusters of fingerprints within a single audio file as disclosed in the subject claim. Wells, *et al.* is silent regarding a parameter (D) describing how evenly spread clusters are, temporally, throughout an audio file.

In addition claim 24 recites *selecting the set of clusters from which to generate the audio thumbnail based upon at least parameter (D).* As discussed above, Petkovic, *et al.* and Wells, *et al.* fail to teach a parameter describing how temporally evenly spread clusters are throughout an audio file. As such, the cited references also fail to teach selecting the set of clusters from which to generate the audio thumbnail based upon this parameter.

In view of at least the foregoing discussion, applicants' representative respectfully submits that Petkovic, *et al.* and Wells, *et al.*, alone or in combination fail to teach or suggest all limitations of applicant's invention as recited in independent claims 1 and 18 (and claims 5, 9, 11, 12, 23, 24 and 34 that respectfully depend there from), and thus fails to make obvious the subject claimed invention. Accordingly, withdrawal of this rejection is respectfully requested.

III. Rejection of Claims 6, 13, 15 and 25-30 Under 35 U.S.C. §103(a)

Claims 6, 13, 15 and 25-30 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Petkovic, *et al.* (US 6,185,527) in view of Nichogi, *et al.* (US 2003/0021472 A1). It is respectfully submitted that this rejection should be withdrawn for at least the following reasons. Petkovic, *et al.* and Nichogi, *et al.*, alone or in combination, do not teach each and every element of applicants' invention as recited in the subject claims.

Claim 6, 13, 15 and 25-30 depend from independent claims 1 and 18 respectively. As noted *supra*, Petkovic, *et al.* does not teach or suggest each and every element of the subject invention as recited in this independent claim, and Nichogi, *et al.* fails to make up for the aforementioned deficiencies of Petkovic, *et al.* Nichogi, *et al.* discloses a system for compressing and decompressing an image. The cited reference is silent regarding audio fingerprints, sets of clusters of fingerprints, or creating a thumbnail from a set of clusters that has at least two gaps between fingerprints. Therefore, Petkovic, *et al.* and Nichogi, *et al.* fail to teach or suggest *producing one or more sets of clusters of fingerprints based upon all fingerprints within a set of clusters meeting a similarity threshold; and creating a thumbnail from a set of clusters that has at least two gaps between fingerprints, wherein a gap is a temporal space between two adjacent fingerprints that exceeds a predetermined threshold when fingerprints within a set of clusters are placed in sequential temporal order.*

Moreover, claim 13 recites *the heuristic component selects the set of clusters from which to generate the audio thumbnail based upon at least one of a mean spectral quality value determined for the set of clusters or a cluster spread quality value determined for the set of clusters.* The cited references fails disclose a set of clusters, let alone a cluster spread quality value for a set of clusters.

Additionally, claim 28 recites *a mean spectral flatness for a set, and a parameter D, are combined to determine a best cluster set from among a plurality of cluster sets.* As conceded in the Office Action with respect to claim 23, Petkovic, *et al.* fails to teach a parameter describing how temporally evenly spread clusters are throughout an audio file. Nichogi, *et al.* is also silent regarding this parameter. As such, the cited references also fail to teach selecting the set of clusters from which to generate the audio thumbnail based upon this parameter.

In view of at least the foregoing discussion, applicants' representative respectfully submits that Petkovic, *et al.* and Nichogi, *et al.*, alone or in combination fail to teach or suggest

all limitations of applicant's invention as recited in independent claims 1 and 18 (and claims 6, 13, 15 and 25-30 that respectfully depend there from), and thus fails to make obvious the subject claimed invention. Accordingly, withdrawal of this rejection is respectfully requested.

IV. Rejection of Claim 31 Under 35 U.S.C. §103(a)

Claim 31 stands rejected under 35 U.S.C. §103(a) as being unpatentable over Petkovic, *et al.* (US 6,185,527) in view of Foote (US 6,542,869). It is respectfully submitted that this rejection should be withdrawn for at least the following reasons. Petkovic, *et al.* and Foote, alone or in combination, do not teach each and every element of applicants' invention as recited in the subject claims.

Claim 31 depends from independent claim 18. As noted *supra*, Petkovic, *et al.* does not teach or suggest each and every element of the subject invention as recited in this independent claim, and Foote fails to make up for the aforementioned deficiencies of Petkovic, *et al.* Foote discloses a system for identifying points of significant change in an audio file that indicate different structural sections of the audio file, such as between music and speech, or differing beats. The cited reference discloses producing a two dimensional matrix that shows similarity between frames of the audio file. The reference also discloses that repetition will appear in the matrix as diagonal stripes. However, Foote fails to disclose that fingerprints that are similar are placed into a set of clusters and that a set of clusters that contains at least two gaps of a predetermined threshold is employed to generate an audio thumbnail. Foote relies upon a user being able to visually translate the matrix into structure. The subject claim provides an automated method for identifying a repeating chorus or instrumental section based upon predetermined criteria, such as sets of cluster, number of clusters, and gaps between clusters. Therefore, Petkovic, *et al.* and Foote fail to teach or suggest ***producing one or more sets of clusters of fingerprints based upon all fingerprints within a set of clusters meeting a similarity threshold; and creating a thumbnail from a set of clusters that has at least two gaps between fingerprints, wherein a gap is a temporal space between two adjacent fingerprints that exceeds a predetermined threshold when fingerprints within a set of clusters are placed in sequential temporal order.***

As such, applicants' representative respectfully submits that Petkovic, *et al.* and Foote, alone or in combination fail to teach or suggest all limitations of applicant's invention as recited

in independent claim 18 (and claim 31 that respectfully depend there from), and thus fails to make obvious the subject claimed invention. Accordingly, withdrawal of this rejection is respectfully requested.

V. Rejection of Claim 32 Under 35 U.S.C. §103(a)

Claim 32 stands rejected under 35 U.S.C. §103(a) as being unpatentable over Petkovic, *et al.* (US 6,185,527) in view of Nichogi, *et al.* (US 2003/0021472 A1) and in further view of Wells, *et al.* (US 2003/0086341 A1). It is respectfully submitted that this rejection should be withdrawn for at least the following reasons. Petkovic, *et al.*, Nichogi, *et al.*, and Wells, *et al.*, alone or in combination, do not teach each and every element of applicants' invention as recited in the subject claims.

Claim 32 depends from independent claim 18. As noted *supra*, Petkovic, *et al.*, Nichogi, *et al.*, and Wells, *et al.* do not teach or suggest ***producing one or more sets of clusters of fingerprints based upon all fingerprints within a set of clusters meeting a similarity threshold; and creating a thumbnail from a set of clusters that has at least two gaps between fingerprints, wherein a gap is a temporal space between two adjacent fingerprints that exceeds a predetermined threshold when fingerprints within a set of clusters are placed in sequential temporal order*** as recited in independent claim 18. Hence, this rejection should be withdrawn.

VI. Rejection of Claims 33, 35 and 36 Under 35 U.S.C. §103(a)

Claims 33, 35 and 36 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Petkovic, *et al.* (US 6,185,527) in view of Kanevsky, *et al.* (US 6,434,520). It is respectfully submitted that this rejection should be withdrawn for at least the following reasons. Petkovic, *et al.* and Kanevsky, *et al.*, alone or in combination, do not teach each and every element of applicants' invention as recited in the subject claims.

Claim 31 depends from independent claim 18. As noted *supra*, Petkovic, *et al.* does not teach or suggest each and every element of the subject invention as recited in this independent claim, and Kanevsky, *et al.* fails to make up for the aforementioned deficiencies of Petkovic, *et al.* Kanevsky, *et al.* discloses a system for archiving audio files by employing natural language processing to convert audio into text to understand the content of the audio file. The system segments the audio file into segments representing music, speech, and background noise and

then translates the speech segments into text. The text is then classified and the audio file archived according to the classification. The cited reference is focused on identifying speech within the audio file to perform natural language processing in order to perform a textual classification of the audio file. Kanevsky, *et al.* is silent regarding placing fingerprints that are similar into a set of clusters and selecting a set of clusters that contains at least two gaps of a predetermined threshold to generate an audio thumbnail. Therefore, Petkovic, *et al.* and Kanevsky, *et al.* fail to teach or suggest *producing one or more sets of clusters of fingerprints based upon all fingerprints within a set of clusters meeting a similarity threshold; and creating a thumbnail from a set of clusters that has at least two gaps between fingerprints, wherein a gap is a temporal space between two adjacent fingerprints that exceeds a predetermined threshold when fingerprints within a set of clusters are placed in sequential temporal order.*

As such, applicants' representative respectfully submits that Petkovic, *et al.* and Kanevsky, *et al.*, alone or in combination fail to teach or suggest all limitations of applicant's invention as recited in independent claim 18 (and claim 31 that respectfully depend there from), and thus fails to make obvious the subject claimed invention. Accordingly, withdrawal of this rejection is respectfully requested.

VII. Rejection of Claim 37 Under 35 U.S.C. §103(a)

Claim 37 stands rejected under 35 U.S.C. §103(a) as being unpatentable over Petkovic, *et al.* (US 6,185,527) in view of Kanevsky, *et al.* (US 6,434,520) and in further view of Wells, *et al.* (US 2003/0086341 A1).. It is respectfully submitted that this rejection should be withdrawn for at least the following reasons. Petkovic, *et al.*, Kanevsky, *et al.*, and Wells, *et al.*, alone or in combination, do not teach each and every element of applicants' invention as recited in the subject claims.

Claim 37 depends from independent claim 18. As noted *supra*, Petkovic, *et al.*, Kanevsky, *et al.*, and Wells, *et al.* do not teach or suggest *producing one or more sets of clusters of fingerprints based upon all fingerprints within a set of clusters meeting a similarity threshold; and creating a thumbnail from a set of clusters that has at least two gaps between fingerprints, wherein a gap is a temporal space between two adjacent fingerprints that exceeds a predetermined threshold when fingerprints within a set of clusters are placed in sequential*

temporal order as recited in independent claim 18. Therefore, this rejection should be withdrawn.

CONCLUSION

The present application is believed to be in condition for allowance in view of the above comments and amendments. A prompt action to such end is earnestly solicited.

In the event any fees are due in connection with this document, the Commissioner is authorized to charge those fees to Deposit Account No. 50-1063 [MSFTP561US].

Should the Examiner believe a telephone interview would be helpful to expedite favorable prosecution, the Examiner is invited to contact applicants' undersigned representative at the telephone number below.

Respectfully submitted,

AMIN, TUROC & CALVIN, LLP

/Himanshu S. Amin/

Himanshu S. Amin

Reg. No. 40,894

AMIN, TUROC & CALVIN, LLP
24TH Floor, National City Center
1900 E. 9TH Street
Cleveland, Ohio 44114
Telephone (216) 696-8730
Facsimile (216) 696-8731